

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 14:42:23 ; Search time 25 Seconds

(without alignments)
658.374 Million cell updates/sec

Title: US-09-745-506-37

Perfect score: 1799
Sequence: 1 MDKALSLSDPFLSPFAE.....LENKINILSLDRPIQVY 350

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1799	100.0	350	NP3L_HUMAN	09gnt8 homo sapien
2	1575	87.5	350	NE3L_MOUSE	09eq80 mus musculu
3	578	32.1	292	A36A_DROME	09nk57 drosophila
4	453.5	25.2	278	YJ3C_SCHPO	09ak04 schizosacch
5	384.5	21.4	288	NIF3_YEAST	P53081 saccharomyc
6	297	16.5	262	YK04_CLOPE	08xv99 clostridium
7	285.5	15.9	373	YQRO_BACSU	P54472 bacillus su
8	278	15.5	366	YF59_STAAM	099t17 staphylococ
9	264	14.7	372	YD03_CLOAB	097110 clostridium
10	263.5	14.6	372	YF89_BACHD	09kd39 bacillus ha
11	256	14.2	373	YF89_LISIN	092p98 listeria in
12	240	13.3	373	YF52_LISMO	P53434 listeria mo
13	225	12.5	241	Y705_CAMUE	09pkr2 campylobact
14	216.5	12.0	379	YK30_MYCTU	010514 mycobacteri
15	210	11.7	262	Y931_STRPY	09a049 streptococ
16	209.5	11.6	263	YK16_ANASP	08y875 anaerobac
17	208	11.6	285	YK01_STRCO	091012 streptomyce
18	203.5	11.3	265	YK09_STRPN	097k80 streptococ
19	199.5	11.1	385	YK30_MYCLE	069411 mycobacteri
20	198	11.0	262	Y351_PNEPA	09p60 ureaplasma
21	193	10.7	243	Y959_HELPY	09zxf2 helicobacte
22	191	10.6	244	Y927_METJA	058337 methanococ
23	187.5	10.4	257	YK73_LACIA	09cgm3 lactococcus
24	181	10.1	251	Y456_MYCPV	025613 helicobacte
25	148	8.2	251	YB61_SALTY	09bq40 mycoplasma
26	133	7.4	247	YB61_SALTY	08xv77 salmonella
27	129	7.2	247	YB61_ECOLI	P75743 escherichia
28	117	6.5	248	Y468_BORBU	051424 borrelia bu
29	112	6.2	252	YK93_VIRBC	09k67 vibrio chol
30	111	6.2	251	Y183_PABMT	09cp72 pasteurella
31	110	6.1	248	Y731_RALSO	08xv40 ralslonia s
32	107	5.9	250	Y627_PYROCO	058361 pyrococcus
33	106	5.9	252	Y855_PSEAE	09hvx2 pseudomonas

34	102.5	5.7	1153	1	A3D1_HUMAN	014617 homo sapien
35	102	5.7	249	1	Y382_NEIMA	09jag6 neisseria m
36	101	5.6	674	1	XPF2_HUMAN	043695 homo sapien
37	100	5.6	249	1	YK54_NEIMB	09jx99 neisseria m
38	97.5	5.4	245	1	YH77_ARCFU	028497 archaeoglob
39	97.5	5.4	502	1	GSPE_PSEAE	000512 pseudomonas
40	97	5.4	250	1	YE24_PYRAB	09uyt3 pyrococcus
41	97	5.4	253	1	YH66_HALNI	09h80 halobacteri
42	97	5.4	263	1	Y110_DEIRA	09t741 delnecoccus
43	97	5.4	488	1	DNAB_HELPY	025916 helicobacte
44	96.5	5.4	247	1	Y291_BUCAP	08k94 buchnera ap
45	96.5	5.4	247	1	YQ97_YERPE	08d91 yersinia pe

ALIGNMENTS

RESULT 1	ID	NP3L_HUMAN	STANDARD:	PRT:	350 AA.
AC	09GZT8; 09H2D2; 09HC18;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	NIF3-like protein 1 (Amyotrophic lateral sclerosis 2 chromosomal				
DE	region candidate gene protein 1) (Miy018 protein) (MDS015).				
GN	NIF3L1 OR ALS2CR1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.				
RX	MEDLINE=20573864; PubMed=1124544;				
RA	Tascon S., Uedelhoven J., Dixkens C., Nayeria K., Engel W.,				
RA	Burfeind P.;				
RT	Isolation and characterization of a novel human gene, NIF3L1, and its				
RT	mouse ortholog, Mif3l1, highly conserved from bacteria to mammals.;				
RL	Cytogenet. Cell Genet. 90:330-336(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21100893; PubMed=11161814;				
RA	Hadano S., Yanagisawa Y., Skaug J., Fichter K., Nasir J.,				
RA	Harindale D., Koop B.F., Scherer S.W., Nicholson D.W., Rouleau G.A.,				
RA	Ikebe J.-E., Hayden M.R.;				
RT	Cloning and characterization of three novel genes, ALS2CR1, ALS2CR2,				
RT	and ALS2CR3, in the juvenile amyotrophic lateral sclerosis (ALS2)				
RT	critical region at chromosome 2q33-q34: candidate genes for ALS2.;				
RL	Genomics 71:200-213(2001).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RA	Meo Y.M., Xie Y., Huang X.Y., Ying K., Dai J.L.;				
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Huang C., Qian B., Tu Y., Gu W., Wang Y., Han Z., Chen Z.;				
RT	"Novel genes expressed in hematopoietic stem/progenitor cells from				
RT	myelodysplastic syndromes patient.;"				
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	TISSUE=Ovarian carcinoma;				
RC	TISSUE=Ovarian carcinoma;				
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,				
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,				
RA	Wagatsuna M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,				
RA	Takahashi M., Chida Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,				
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,				
RA	Yamamoto Y., Makatsutsu A., Nakamura Y., Nagahara K., Masuno Y.,				
RA	Ninomiyu K., Iwayanagi T.;				
RT	"NEO human cDNA sequencing project.;"				
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.				
RN	[6]				

RP SEQUENCE FROM N.A.
 RC TISSUE-Skin;
 RX MEDLINE=23388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Hosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE UPR0135 (NIF3) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF283538; AAC4846.1; ALT_INIT.
 DR EMBL; AB038949; BAB32499.1; -
 DR EMBL; AF060513; AAC4313.1; -
 DR EMBL; AF182416; AAC4952.1; -
 DR EMBL; AK023378; BAB4551.1; -
 DR EMBL; BC007654; AAH07654.1; ALT_INIT.
 DR Genew; HGNC:13390; NIF3L.
 DR Interpro: IPR002678; DUF34.
 DR Pfam: PF01784; NIF3.1.
 DR TIGRFAMS; TIGR00486; TIGR00486.1.
 DR CONFLICT; 77 82 TWNTWK -> NLEHNGR (IN REF. 4).
 SQ SEQUENCE 350 AA; 38983 MM; 81FLASAD35B25ED) CRC64;
 Query Match 100.0%; Score 1799; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 1e-135;
 Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDLALLSLNDFASLSFAESMDVNGLLVPSPPHTVNTLFLTNDLTEEVEVLQKRD 60
 DB 1 MDLALLSLNDFASLSFAESMDVNGLLVPSPPHTVNTLFLTNDLTEEVEVLQKRD 60
 QY 61 LILSYHPIRFPMKRITWNTWKERLYTALERNVCITSPHAYDAAPGVNNWLAKGGA 120
 DB 61 LILSYHPIRFPMKRITWNTWKERLYTALERNVCITSPHAYDAAPGVNNWLAKGGA 120
 QY 61 LILSYHPIRFPMKRITWNTWKERLYTALERNVCITSPHAYDAAPGVNNWLAKGGA 120
 DB 61 LILSYHPIRFPMKRITWNTWKERLYTALERNVCITSPHAYDAAPGVNNWLAKGGA 120
 QY 121 CTSRPIHPSKAPNPTGEGNHRFVNVTYDLDKMSAVKGIQSVTSFSARIGNEQT 180
 DB 121 CTSRPIHPSKAPNPTGEGNHRFVNVTYDLDKMSAVKGIQSVTSFSARIGNEQT 180
 QY 121 CTSRPIHPSKAPNPTGEGNHRFVNVTYDLDKMSAVKGIQSVTSFSARIGNEQT 180
 DB 121 CTSRPIHPSKAPNPTGEGNHRFVNVTYDLDKMSAVKGIQSVTSFSARIGNEQT 180
 QY 181 RININCTOKALMOVDFLSRNKOLYOKTEILSLKPLLLHTGMGRICLTIDSVSLATMID 240
 DB 181 RININCTOKALMOVDFLSRNKOLYOKTEILSLKPLLLHTGMGRICLTIDSVSLATMID 240
 QY 181 RININCTOKALMOVDFLSRNKOLYOKTEILSLKPLLLHTGMGRICLTIDSVSLATMID 240
 DB 181 RININCTOKALMOVDFLSRNKOLYOKTEILSLKPLLLHTGMGRICLTIDSVSLATMID 240
 QY 241 RIKRHLKLSHRLALGVRITLESQYKVALCAGSSSVLQGVADLYITGEMSHDITDA 300
 DB 241 RIKRHLKLSHRLALGVRITLESQYKVALCAGSSSVLQGVADLYITGEMSHDITDA 300
 QY 241 RIKRHLKLSHRLALGVRITLESQYKVALCAGSSSVLQGVADLYITGEMSHDITDA 300
 DB 241 RIKRHLKLSHRLALGVRITLESQYKVALCAGSSSVLQGVADLYITGEMSHDITDA 300
 QY 301 ASGGINVLCEHSNTERGFLSDLRMDLSHLENKINILSETDRDPLQV 350
 DB 301 ASGGINVLCEHSNTERGFLSDLRMDLSHLENKINILSETDRDPLQV 350

DB 301 ASGGINVLCEHSNTERGFLSDLRMDLSHLENKINILSETDRDPLQV 350
 RESULT 2
 ID NIF3L_MOUSE STANDARD; PRT; 350 AA.
 AC O9R080: O9D098;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NIF3L-like protein 1.
 GN NIF3L.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RX MEDLINE=20573864; PubMed=11124544;
 RA Tascou S., Uedelhoven J., Dixkens C., Nayeria K., Engel W.,
 RA Bufeind P.,
 RT "Isolation and characterization of a novel human gene, NIF3L, and its
 RT mouse ortholog, Nif3l, highly conserved from bacteria to mammals."
 RL Cytogenet. Cell Genet. 90:330-336(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Flatschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gunsberg S., Hill D., Hofmann M., Hume D.A., Kanalya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts F.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- SIMILARITY: BELONGS TO THE UPR0135 (NIF3) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF284439; AAC45961.1; ALT_INIT.
 DR EMBL; AK011670; BAB27769.1; ALT_INIT.
 DR MGD; MGI:1929485; NIF3L1.
 DR Interpro: IPR002678; DUF34.
 DR TIGRFAMS; TIGR00486; TIGR00486.1.
 DR CONFLICT; 164 164 G -> S (IN REF. 2).
 FT CONFLICT; 178 178 E -> K (IN REF. 2).
 FT CONFLICT; 191 191 L -> F (IN REF. 2).
 FT CONFLICT; 195 195 L -> Q (IN REF. 2).
 FT CONFLICT; 198 198 L -> F (IN REF. 2).
 FT CONFLICT; 208 208 T -> I (IN REF. 2).
 SQ SEQUENCE 350 AA; 38828 MM; B0FA71503BCF7086 CRC64;

Query Match 87.5%; Score 1575; DB 1; Length 350;
 Best Local Similarity 86.3%; Pred. No. 6,5e-118;
 Matches 302; Conservative 24; Mismatches 24; Indels 0; Gaps 0;

QY 1 MDKALLSLNDFAISFAESMDVNLVPPSPHTVNTLTNDLTFEEMEVLOKKAD 60
 DB 1 MDKALLSLNDFAISFAESMDVNLVPPSPHTVNTLTNDLTFEEMEVLOKKAD 60
 QY 61 LILSYHPPIFRPKKRTWTMTWKRLVIRALENRVGIYSPHTADAPQGNVNLGLGA 120
 DB 61 LILSYHPPIFRPKKRTWTMTWKRLVIRALENRVGIYSPHTADAPQGNVNLGLGA 120
 QY 61 FILSYHPPIFRPKKRTWTMTWKRLVIRALENRVGIYSPHTADAPQGNVNLGLGT 120
 DB 121 CTSRPIHPSKAPNYPTGEGHNVFENNYTQDLKVMASVAGIDGVSPTSARTGNEBOT 180
 QY 121 CTSRPIHPSKAPNYPTGEGHNVFENNYTQDLKVMASVAGIDGVSPTSARTGNEBOT 180
 DB 121 CTSRPIHPSKAPNYPTGEGHNVFENNYTQDLKVMASVAGIDGVSPTSARTGNEBOT 180
 QY 181 RINLNCOTKALMVDVFLSRNKKLYOKTEILSLKPLLLHTGRLCTIDESVSLATMID 240
 DB 181 RINLNCOTKALMVDVFLSRNKKLYOKTEILSLKPLLLHTGRLCTIDESVSLATMID 240
 QY 241 RIKRHKLSHIRALAGVGRLESQVAVVACAGSSSVYAGVADLYTGEKSHHTLDA 300
 DB 241 RIKRHKLSHIRALAGVGRLESQVAVVACAGSSSVYAGVADLYTGEKSHHTLDA 300
 QY 301 ASGGINVILCEHSNTERGFLSDRLMDLSLEKNINILSETDRDPLQVY 350
 DB 301 ASGGINVILCEHSNTERGFLSDRLMDLSLEKNINILSETDRDPLQVY 350

RESULT 3
 A36A.DROME STANDARD; PRT; 292 AA.

ID A36A.DROME
 AC 09NK57; 09VIL4;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Anon-35F/36A protein.
 GN ANON-35F/36A OR BG:DS02740.16 OR CG4278.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Phylotribe; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazey R.G., Davis T.,
 RA Doyle C., Galle R.F., George R.A., Harris N.L., Hartzell G.,
 RA Harvey D., Hong L., Houston K.A., Hoskins R.A., Johnson G., Martin C.,
 RA Moshrefi A., Palazzo M., Reese M.G., Spradling A.C., Tsang G.,
 RA Wan K.H., Whitelaw K., Celniker S.E., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of
 RT Drosophila melanogaster: the Adh region.";
 RL Genetics 153:179-219(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agayari A., An H.-J., Andrews-Planck C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova K., Botchan M.R., Bouck J., Broksstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
 RA Jastai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Malsbina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Stupak M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE003415; AAF45006.1; -
 CC EMBL: AE003650; AAF53525.1; AL7-INT.
 CC DR FLYBASE: FBgn0014092; CG4278.
 CC DR InterPro: IPR002678; DUF34.
 CC DR Pfam: PF01784; NIF3.1.
 CC DR TIGRFBMS: TIGR00486; TIGR00486.1.
 CC DR SEQUENCE 292 AA; 3263 MW; 775D27E7721150CC CRC64;
 SO SEQUENCE

Query Match 32.1%; Score 578; DB 1; Length 292;
 Best Local Similarity 33.9%; Pred. No. 9.6e-39;
 Matches 118; Conservative 61; Mismatches 81; Indels 88; Gaps 4;

QY 3 LKALLSLNDFAISFAESMDVNLVPPSPHTVNTLTNDLTFEEMEVLOKKADLI 62
 DB 13 LAAVVKELNFAPTSMKEKNDVNLVPPSPHTVNTLTNDLTFEEMEVLOKKADLI 72
 QY 63 LSHYHPPIFRPKKRTWTMTWKRLVIRALENRVGIYSPHTADAPQGNVNLGLGA 122
 DB 73 LSHYHPPIFRPKKRTWTMTWKRLVIRALENRVGIYSPHTADAPQGNVNLGLGA 122
 QY 123 SRPIHPSKAPNYPTGEGHNVFENNYTQDLKVMASVAGIDGVSPTSARTGNEBOTRI 182
 DB 133 IRPLEPE-----LGAPG----- 145
 QY 183 NLNCTOKALMVDVFLSRNKKLYOKTEILSLKPLLLHTGRLCTIDESVSLATMIDRI 242
 DB 146 -----TSGRY--IETKMSLSQVVEST 165
 QY 243 KRLKLSHIRALAGVGRLESQVAVVACAGSSSVYAGVADLYTGEKSHHTLDAAS 302
 DB 166 OKRIKRS-VHVAALVAGTPPTLTOSVIGICGASASLKGIOADLITGESHHEVLETH 224
 QY 303 OGINVILCEHSNTERGFLSDRLMDLSLEKNINILSETDRDPLQVY 350
 DB 225 NNTVILCNHSNTERGFLHEFCETLAKSLNECLVAFVSVDKDPVLYV 272

RESULT 4
 YJ3C_SCHPO

```

ID K13C_SCHPO STANDARD: PRT: 278 AA.
AC 094404;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical protein C126.12 in chromosome III.
CN SPC126.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brookes K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Moore P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K., Stevens K.,
RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hiltner H.,
RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
RA Egger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleaux V., Mottier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1 SIMILARITY: BELONGS TO THE UPR0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, AL034490; CAA22481.1; -
CC PIR, T40916; T40916.
DR GenDB_Spombe: SPC126.12; -
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIF3; 1.
DR TIGRfams: TIGR00486; TIGR00486; 1.
KM Hypothetical protein.
SQ SEQUENCE 278 AA: 29850 MW: A3314E79092D677C CRC64;

```

```

Query Match 25.2%; Score 453.5; DB 1; Length 278;
Best Local Similarity 32.7%; Pred. No. 7.2e-29;
Matches 116; Conservative 53; Mismatches 93; Indels 93; Gaps 11;

```

```

QY 4 KALLSS-----LNDPASISFAESPMNVGLIVE-PSPPHVNTLFTNLTETVEYEV 57
DB 3 KANISKLKVVESITYNKRLADSDNNTGLLEAPPRRNASSVLLTIDITKVAEALSN 62
QY 58 K-ADLLSYHPIFRPMKRITWNTWKERLVIRALENRVGIYSPHTAYDAAPGVNNW 116
DB 63 KLVSSIVAVHPIIFGLKAIWEDPQORSLLKLAEGIHVYSPHTAVDAADVNDMLAQ 122

```

```

QY 117 GLGACTSPHPKAPNPPTGNNHRENNVNTODDLKVMASVKGIDGVSTFSARTGN 176
DB 123 GIAG-----GRNNIK-SVPTQ----- 138
QY 177 EEOFRINLCYOKALMOYVDFLSRNKOLYQKTEILSLEKPLLHTGMGLCTLDESYSIA 236
DB 139 -----QNSVMAFAE-----GYRICELKIPPTLR 162
QY 237 TMDIRKHLKLSHRIALGVGRFLSQVKVALCAGSGSSVLOGVEADLYTGMSSHHD 296
DB 163 ELVGRKELTGLGYQVOCAPNG--LDHSISKVSICAGSGSGSVNMNTADLTFTGELSHHQ 220
QY 297 TLDAASGGINVILCEHSTNTERGFLSDL--RMDLS-HLEN-KINILISFTRDPL 347
DB 221 VLAAMAKGISVILGHSNTERGYLKDVCQKRLASSFHEKGVADVAVYSSMDADPL 275

```

RESULT 5

```

NIF3_YEAST STANDARD: PRT: 288 AA.
ID NIF3_YEAST
AC P53081;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ngi1-interacting factor 3.
GN NIF3 OR YGL22IC.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Bruechner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII.";
RL Yeast 13:1077-1090(1997).
CC -1 SIMILARITY: BELONGS TO THE UPR0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, 272743; CA96937.1; -
CC PIR, S64243; S64243.
DR SGD: S0003189; NIF3.
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIF3; 1.
DR TIGRfams: TIGR00486; TIGR00486; 1.
SQ SEQUENCE 288 AA: 31888 MW: B6ABE48BAF4776A CRC64;

```

```

Query Match 21.4%; Score 384.5; DB 1; Length 288;
Best Local Similarity 28.6%; Pred. No. 2.3e-23;
Matches 104; Conservative 59; Mismatches 98; Indels 103; Gaps 11;

```

```

QY 3 LKALLSSLNDPASISFAE-SWMDNGLVYEPSPPHVNT-----LFTNLTETVEYEV 54
DB 10 LDKLVRSITTKFEPQRYAKSDWNTGLLIDCSAQVYTTADANAKTKVLLTVLDTVSAD 69
QY 55 LOKKADLLSYHPIFRPMKRITWNT-WKERLVIRALENRVGIYSPHTAYDAAPGVNNW 113
DB 70 VDANCNVIAVHPIFRPMKRNRPHTNPQHEATKLIYGSVYCPHTAVDAARGVNDW 129
QY 114 LAKGLGACTSPHPSKAPNPPTGNNHRENNVNTODDLKVMASVKGIDGVST-SFSA 172
DB 130 LVKGLN-----NENNAKSYAL 146
QY 173 RTGNEEOFRINLCYOKALMOYVDFLSRNKOLYQKTEILSLEKPLLHTGMGLCTLDES 232

```

DB 147 ETVSGE-----TDDL-----IGXGRFEFNKD 168
QY 233 VSLATMDIRKRLKSHIRLALGVGRILESOYKV--VALCAGSGSSVLYGV--EADLYL 288
DB 169 ISLEQIVKVKRRLRPVYVAVSLAAPSANNOLKIKRVAACAGSGGVFQQLKEDVDLY 228
QY 289 TGEASHHDTLDAASQGINVILCEHSNTERGFLSDRLDML---DSHLENKINILSETDRD 345
DB 229 TGEASHHEVLKMKEMKRYIVCNHNTERGFLDQVAKGLLODEGH-----EYVASKMDCD 283
QY 346 PLQV 349
DB 284 PLTV 287
RESULT 6
YK04_CLOPE STANDARD: PRT: 262 AA.
ID YK04_CLOPE
AC 08XIV9;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical protein CPE2004.
GN CPE2004.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shima T., Ogasawara N., Hattori M., Kihara S., Hayashi A.,
RT Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1- SIMILARITY: BELONGS TO THE UP0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP003192; BAB81710.1; -
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIF3.1
DR TIGRPFAM: TIGR00486; TIGR00486.1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 262 AA: 29127 MM; 8CC96C79178756D7 CRC64;
Query Match 16.5%; Score 297; DB 1; Length 262;
Best local similarity 23.5%; Pred. No. 1, 8e-16;
Matches 83; Conservative 68; Mismatches 108; Indels 34; Gaps 7;
QY 1 MDKALLSLNDFAISFAESMDNVLVEPSPHVTNLTNDLLEVEEYLOKKAD 60
DB 1 MKINDIINIEDIAPNLEKEGFNDVGMVGDREKN-ITRILLALDCTEEYIKRAKMGAE 59
QY 61 LILSYHPITFRPMKRTITMTWKERIVIRALENVGIIYSPHTADAPQGVNNLALGLCA 120
DB 60 LILTHPLPLFRKPSITTTDPLLRKTIISLIKNDINISAHNTMDSVKGLNDLVEILG- 118
QY 121 CTSRPIHPSKAPNYPTGHNHREVFNNYTOTDDIKVASAKGIDGVSVTSFARTGMEQ 180
DB 119 -----FN-----KGI----- 123
QY 181 RINLNCOTKALMOVDFLSRNKOLOKTEILSLERPLLTGKRLCTLD3SVSLATMID 240

DB 124 -----IMD-----KSPVDSBAGIGRVELTKEMTVLEIIN 153
QY 241 RIRKRLKSHIRLALGVGRILESOYKVVALCAGSGSSV---QGVENDLYLTGEMSHHD 297
DB 154 LIRKSSIGVKNLRYA---GDLENYIKRIALVNSGODFGDARKLGADLIITDPTTHFV 209
QY 298 LDAASQGINVILCEHSNTERGFLSDRLDMLSDHLENKINILSETDPLQV 350
DB 210 SDYKEMGLNILDIGHFENSEMPVILIKVSEKVERLSDVEITVSKAKDPEPI 262
RESULT 7
YK04_BACSU STANDARD: PRT: 373 AA.
ID YK04_BACSU
AC P54472; 032024;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical protein ygfO.
GN ygfO.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.,
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes."
RL Microbiology 142:3103-3111(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertorello M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borstis R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Chou S.K., Codani J.J., Conerton J.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Deville K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Ertling K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritze C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
RA Kobayashi Y., Koeter P., Konigstein G., Krogh S., Kunano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Meisel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portelle D., Portwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scotlone F.,
RA Sekiuchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorekchi A., Taccani E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takemoshi A., Tanaka T., Terstra P., Tognoni A.,
RA Tosi V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T.,
RA Winters P., Wiput A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1- SIMILARITY: BELONGS TO THE UP0135 (NIF3) FAMILY.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 270.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; D84432; BAI2492.1; ALT_FRAME.
 DR EMBL; 299116; CAB14447.1; -
 DR PIR; A69954; A69954.
 DR Subtilisin; Bg1161; ygfO.
 DR InterPro: IPR002678; DUF34.
 DR Pfam: PF01784; NIF3.1
 DR TricRfams; TricR00486; TricR00486; 1.
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 265 266 KD -> FC (IN REF. 1).
 SQ SEQUENCE 373 AA; 40920 MW; E2320C1268DBE9B CRC64;

Query Match 15.9%; Score 285.5; DB 1; Length 373;
 Best Local Similarity 26.1%; Pred. No. 2.4e-15;
 Matches 99; Conservative 79; Mismatches 143; Indels 59; Gaps 17;

QY 6 LLSLSLNDFAISLFAESMDNVGL---LVEPSPTHTVTLFLNDLLEVEEVLQKADL 61
 DB 10 ILQFEGFSPKAYAVEGDKIGLIGTLNKP---IKNVVTLDYLEGVIDEALKEVDL 64
 QY 62 ILSTHPTFRPMKRTITNTWKERLYRALNRYGYSPTAYDAPOGVNMLAKGLGAC 121
 DB 65 ILAHPPTFRSLKSTDPAGRLIEKCLHDIAVYAHNTMDVADGVNLDLAEALSL 124
 QY 122 TSRIHPKAPNY-----PTE-----GN-----HRYEFVNTYTDLDKYS 157
 DB 125 ETEVL---APTYDPLKLAIVYVPEEYORALAGAGHIGY---SHCATSSSEIG 178
 QY 158 AVKGIDVSVYSFSARTGNEE---QTRINLCTOKALMQVDFLSRNKQLYOKT---ELLIS 212
 DB 179 SFKPLDGAK--PTIGEVELELVHEVRLVETVPKSEKAVANAMKSHP--YEEVAYDIYP 235
 QY 213 LEKPLLTGMRGLCTLDESVSLATMDIRIKRHLKLSIRLALGVRLESOYKVALCA 272
 DB 236 VEO--TPAEKGLGRVOTLKNEMTLKEFALFVKDKLDVNCVRM---VG--DADSVKVAVLG 290
 QY 273 GSGSVLQGV---ADLYTGEMSHHDTLDAASOGINYLCEHSTERTGFLSDRLDMS 329
 DB 291 GDGNKTYIHAKRKAGDYVYTGDIYFVAHADMLGLNVVDGQH--YAEIKMEGYTRKILTS 349
 QY 330 HLENK---INITLSETRDP 346
 DB 350 MCNDKFGVNIFVSETDTP 369

RESULT 8
 YF59_STAM
 ID YF59_STAM STANDARD; PRT; 366 AA.
 AC 099T7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein SAV1559/SA1388.
 GN SAV1559 OR SA1388.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=158878, 158879;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC STRAIN-Mu50 / ATCC 700699, and N315;
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian Y.-Q., Ito T.,
 RA Karamori M., Matsumaru H., Murakami H., Hosoyama A.,
 RA Mizutani U. Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,

RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogatawara N., Hayashi K., Hirakawa K.,
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus.";
 RL Lancet 357:1225-1240(2001).
 CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AP003362; BAB57721.1; -
 DR EMBL; AP003362; BAB57721.1; -
 DR PIR; F89936; F89936.
 DR InterPro: IPR002678; DUF34.
 DR Pfam: PF01784; NIF3.1
 DR TricRfams; TricR00486; TricR00486; 1.
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 366 AA; 41075 MW; CA7D5FADFD78FF74 CRC64;
 SQ SEQUENCE 366 AA; 41075 MW; CA7D5FADFD78FF74 CRC64;

Query Match 15.5%; Score 278; DB 1; Length 366;
 Best Local Similarity 26.0%; Pred. No. 9.4e-15;
 Matches 105; Conservative 67; Mismatches 140; Indels 92; Gaps 18;

QY 1 MDLALLSLNDFASLFAESMDNVGLVEPSPTHTVTLFLNDLLEVEEVLQKAD 60
 DB 1 MKIDLMTLLDHHVPESTAESMDNVGLIGDGEVTVGL-TALDCTLEVNEAIEKGYN 59
 QY 61 ILSTHPTFRPMKRTITNTWKERLYRALNRYGYSPTAYDAPOGVNMLAKGLG 119
 DB 60 TTIHHPPTFRSLKSTDPAGRLIEKCLHDIAVYAHNTMDVADGVNLDLAEALSL 117
 QY 120 ACTSRPIPSKAPNYPT-----EGNHRVE----- 144
 DB 118 LKNSIINODVYKQVYIPKDNVGFPKKLSNGLAGCNVEYCFESEGGQFV 177
 QY 145 -NVNVT--ODLKWASAVKGDGVSVTSARTGNEQTRINLCTOKALMQVDFLSNK 202
 DB 178 GEANPTTQIDK---IEDVDEVKI-----EFMIDVYOK-----SRAE 211
 QY 203 QLYOK-----TELSLEKPLLTGMRGLCTLDESVSLATMDIRIKRHLKSLRLA 254
 DB 212 QLIKQHPYETPVDFEIKQSLX--GLGVAAEVDNOMTLEDPAADIKSLNTPSVAF- 268
 QY 255 LGVGRTELOKVVVALCAGSG---SSVLOGVEADLYLTGEMSHHDTLDAASOGINYL 309
 DB 269 --VGES--NQIKIRLAIIGSGIGYEYQAVQCG--ADVEVTDGDKHHDALDAKIHGVLND 323
 QY 310 CEHSTERTGFLSDRLDMSLHLE-NKINIL--SETDQDPLQV 350
 DB 324 INH-YSEYVMEKGLTKTLTMNFMENIKINIDVDEASTINDPROXI 366

RESULT 9
 YD03_CLOAB
 ID YD03_CLOAB STANDARD; PRT; 268 AA.
 AC 097T10;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein CAC1303.
 GN CAC1303.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;

RA MEDLINE-21359325; PubMed-11466286;
RA Noelling J., Breton G., Omeilenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hittl J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE007642; AAK79274.1; -
DR PIR; G97060; G97060.
DR InterPro: IPR002678; DUF34.
DR Pfam; PF01784; NIF3; 1.
DR TIGRPFAM; TIGR00486; TIGR00486; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 268 AA; 29850 MW; AAF886E324861AB2 CRC64;

Query Match 14.7%; Score 264; DB 1; Length 268;
Best Local Similarity 21.3%; Pred. No. 8e-14;
Matches 75; Conservative 63; Mismatches 118; Indels 36; Gaps 6;

QY 1 MDLKLALSSINDPASLFAESMDNVGLVEPSPHTVNTLFLTNDLVEEYMEVLOKKAD 60
DB 3 LAKDLCNIEDPAPSLKEDFDNGLAMGDREA-SYDAIMRALDCTMDYIDAIKKN 61
QY 61 LLSLHPPPLPFPKRTITNTWKERLYRLENRVGISPHYDAAPQGVNMLALGCA 120
DB 62 MITTHPILFKKPKSKITMDTLGKKIKIISNNINYSANTNIDSVKDGINDAVNVLG- 120
QY 121 CTSRPHPSKAPNPYEGNHVREENVNYQDLDKVASAKGIDGVSVTSPSANTGNEQT 180
DB 121 ----- 120
QY 181 RLNLCTOKALQVYDFLSRNKOLYOKTEILSLKPLLLHTGNGRLCTLDESVSATMTID 240
DB 121 -----FDKSSILAKNNKAVKEAGIGRVELEQNMNTLAKELD 156
QY 241 RIKRHUKLSHRLALGVGRTLESQVAVVVALCAGSSSVLQGV-ADLYLGEMSHHOT 297
DB 157 RVKESFKIQLSRVCGDEKRIHS---FAVINGSGGDFEERARRGVDCTITGDTSYHYV 212
QY 298 LPAASGGINVILCEHSNTERGFLSLDRMLDLSHLENK-IN--ITLSETPD 346
DB 213 SPTNEMNINAVIDAGHGTETMPSPVYVMSKLLGALHMGINTPLIVSQNNIDP 264

RESULT 10
ID YD80_BACHD STANDARD; PRT; 372 AA.
AC O9K039;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BHI380.
GN BHI380.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE-20512582; PubMed-11058132;
RA Takai H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horikoshi K.;
RT "Complete genome sequence of the alkaliophilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001511; BAB05099.1; -
DR PIR; D83822; D83822.
DR InterPro: IPR002678; DUF34.
DR Pfam; PF01784; NIF3; 1.
DR TIGRPFAM; TIGR00486; TIGR00486; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 372 AA; 40813 MW; DD79BF6630CE5B CRC64;

Query Match 14.6%; Score 263.5; DB 1; Length 372;
Best Local Similarity 24.9%; Pred. No. 1.4e-13;
Matches 89; Conservative 80; Mismatches 141; Indels 47; Gaps 12;

QY 25 VGLVPSPPHYVNTLFLTNDLVEEYMEVLOKKADLLSYNPPITRPKRTITNTWKER 84
DB 32 IGLTLPK-----IQRLVTLADYVESYIDEVIEGALILAHNPRIPLSSITDPAVY 86
QY 85 LVTRALENRVIGISPHYDAAPQGVNMLALGCACTSPRHPKAPNPYEGNHVRE 144
DB 87 IIRKALKHDLITAAATNLDITGKGVNDLADLGLKIDELVAP-----TTSLYKLAV 141
QY 145 NVNVTODLDKVASVKGIDGVSYSFASRTGNEQT-----RLNLCTOKALQVYDF 197
DB 142 FVPHTH-TQVREALRAGAGHIGNYSCTFNSKGTGFEKPEGTNPFGKQALFEVE 200
QY 198 L-----SRKOLYOKTEILSLKPLL-LH-----TGMGRLCTLDESVSATMT 239
DB 201 LKLTETVTEGQKKKVAAMIKSHPYEPAYDYLPLANEGTLLGLGRIYHESMTLEERA 260
QY 240 DRIKRLKLSHRLALGVGRTLESQVAVVVALCAGSSSVLQGV-EADLYLGEMSHHD 296
DB 261 KQYKKAFFDVPYAV--VG-SLETQIRKAVAVLGDDNKWMAHLRKGAIVYIGVYYNH 316
QY 297 TLDAASGGINVILCEHSNTERGFLSLDRMLDLSHLENK--INIISETD 350
DB 317 PHDALMDGINTVDPGH-NVEKIMKQGVKELKLDKRYDEVVAVSVHTDPETFI 372

RESULT 11
ID YE89_LISTIN STANDARD; PRT; 373 AA.
AC O92B08;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Ltn1489.
GN Ltn1489.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CLIP 11262 / Serovar 6a;
RX MEDLINE-21537279; PubMed-11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,
RA Dommann E., Dominguez-Bernal G., Duchaud E., Durand L., Dussuguet O.,
RA Entlian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurpik G.,
RA Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.,
RT "Comparative genomics of *Listeria species*".
RL Science 294:849-852(2001).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce>
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; AL596168; CAC96720.1; -
DR PIR; AH1618; AH1618.
DR ListList; L1N01489; -
DR InterPro; IPR002678; DUF34.
DR Pfam; PF01784; NIF3.1.
DR TrifRfam; TIGR00486; TIGR00486.1.
KW Hypothetical protein; Complete proteome.
KW SEQUENCE 373 AA; 41436 MW; 4BBC6A93B0B9F540 CRC64;
SO SEQUENCE

Query Match 14.2%; Score 236; DB 1; Length 373;
Best Local Similarity 22.5%; Pred. No.5,4e-13;
Matches 86; Conservative 78; Mismatches 152; Indels 66; Gaps 11.

OY 10 LNFASLSFASPMQVGLVSPSPHYNTLFTNDLBEVMEVLOKKADLLSYHPPI 69
DB 13 MEKIAPKKLAEBGDTGLQVGDLSKKVKKVM-TIDVLEVVDAIEKRVLDLIAHHPEL 71
OY 70 FRPMKRITWNTWKETLVRALENRGVSYPTAYDAAGVNNMLAKGIGACTSRPIHS 129
DB 72 YRPQHIDFTTKQSGMKIKLLKHDTVRAHNTINDIAGGVNDIADLLHIDDTMIEST 131
OY 130 -----KAPNPTPE-----GNHRVEFNVTYDLDKVMKAGIDGVSYTSFSA 172
DB 132 YTEPYCKIAVYVPENELESVRALAVNMGAGIGTYNTE---CPHTTGIGSFKPGTDAN 187
OY 173 RTGNEQGRINL-----NCTQRAIM-----QVDFLSNKKLQYKTEL 211
DB 188 PTIGKEKELTSLPEVKIARFPQYLTEITTKAVKIAHRYEERPAIDVYTLLETQYKE---- 243
OY 212 SLERPLLTHTGGRCLTIDESVSLATMIDRIKRLKLSHRIALGVRGTLSEQVVAALC 271
DB 244 -----GLGRGTLPKRKISWVSLDKLTKTAIAIDNVR---VG-DLRANQKVAIT 289
OY 272 AGSGGSYL--QGVADYLVTGEMSHNDTLDAASAGINVILCENS--NTERGFLSDLRDM 326
DB 290 GGDGKRFTHQAKATGADVFTIGDYVYHHAHDLALNPLTIDAGHNIEKMKYGLKNKME 349
OY 327 LDSHLENKINILSETDDPQL 348
DB 350 QAKIDYEAEFIVSEVNTDPQ 371

RESULT 12
ID YE52_LISMO STANDARD; PRT; 373 AA.
AC P53434;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Lmo1452.
GN LMO1452.
OS *Listeria monocytogenes*.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.

[illegible]


```

RESULT 13
ID Y705_CAMJE STANDARD: PRT: 241 AA.
AC 09PK2:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein cj0705.
GN cj0705.
OS Campylobacter jejuni.
OC Bacteria: Proteobacteria: Epsilonproteobacteria: Campylobacteriales;
OC Campylobacteriaceae: Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MCRC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wen B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jagsels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
CC -1 SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

YM30_MYCTU
ID YM30_MYCTU STANDARD: PRT: 379 AA.
AC 010514;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2230C.
GN RV2230C OR MT2289 OR MYC4427.11C.
OS Mycobacterium tuberculosis.
OC Bacteria: Actinobacteria: Actinomycetales;
OC Corynebacterineae: Mycobacteriaceae: Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC
CC SEQUENCE FROM N.A.
RN RP STRAIN=CDC 1551 / Oshkosh;
RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Esmailova M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikala A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -1 CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
CC IN POSITIONS 254 AND 289.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

